Noname manuscript No.
(will be inserted by the editor)

Detecting Object Motion in Dynamic Settings using Passive RFID

Siddika Parlak · Ivan Marsic

Received: date / Accepted: date

Abstract We explore the problem of long-range object motion detection using passive RFID. Unlike the previous work, we focus on dynamic settings suffering interference caused by humans and multiple tags. We created a realistic setting including these factors and generated a dataset of approximately 14000 detection instances. We observed that actual tag motion and other factors affecting radio signal have different statistical fingerprints; hence, they can be distinguished using statistical methods. Our methodology for object motion detection depends on extracting descriptive features from the received signal strength and classifying them using machine-learning techniques. We report experimental results obtained with several statistical features and classifiers, and we provide guidelines on feature and classifier selection under different constraints and environmental conditions. Experimental results showed that object motion detection, up to 85% accuracy, is achievable even in challenging scenarios. Type of motion, on the other hand, could not be identified with such high accuracy using current passive RFID technology.

1 Introduction

Mobility of a user, or an object, provides valuable information for building context-aware applications, such as assisting the elderly people depending on their activity levels, monitoring medical equipment utilization in hospitals or recognizing performed activities by monitoring use of objects. The object to be tracked is often equipped with a sensor, and sensor type is selected based on the constraints (e.g., cost, object size) and requirements (e.g., precision, latency) of the application.

Siddika Parlak
Rutgers University
Tel.: +1 732 445 4208
E-mail: parlak@rutgers.edu

Ivan Marsic
Rutgers University
Tel.: +1 732 445 6400
Fax.: 1 732 445 4775
E-mail: marsic@rutgers.edu

Radio-Frequency Identification (RFID) technology offers a non-intrusive, low-cost and privacy-preserving solution for sensing small and inexpensive objects. Compared to passive identification technologies (e.g., barcodes), RFID provides faster and simultaneous scanning of multiple items, longer read-range and non line-of-sight operability, minimizing the need for human effort. Compared to active sensors (e.g., accelerometers and active RFID tags), passive RFID tags do not require maintenance because they operate without batteries. They are also smaller, which makes them convenient for attaching to small objects, and cheaper, which makes them usable at the item level and even disposable. Although computer vision also offers most of these advantages, it raises privacy concerns. Cameras provide a permanent visual record of people and their activities, whereas RFID data includes only limited or no personal information. Moreover, RFID technology is better at detecting small and randomly oriented objects, and is more appropriate for dealing with occlusions than computer vision algorithms [28]. Despite its advantages, long-range passive RFID has received limited attention in activity recognition community due to its performance. Near-field RFID readers (fixed or worn) have been used for robust tag detection [6,17,23,25]; however, they are intrusive for human activities because they require users to remember to wear them, or scan the objects through readers similar to scanning barcodes.

In this work, we explore using passive ultra-high frequency (UHF) RFID for long-range (2-3 m) detection of object motion within a room. We define *object motion* as any movement experienced by the object due to human interaction, such as holding or relocating. Interactions with an object cause a change in its orientation and location, as well as occlusions with hand or body. These resulting effects are also the factors determining the energy reflected by RFID tags. Continuous interactions will cause high and frequent changes in Received Signal Strength Indication (RSSI). These changes have different statistical characteristics compared to the fluctuations caused by environmental changes, such as human movement near the tag. Our methodology for detecting motion depends on processing the RSSI sequence to detect fluctuations due to tag motion.

We started by listing the constraints and challenges introduced by different domains that might affect motion detection performance. The dynamic nature of our target application, recognizing tasks during trauma resuscitation, required using statistical methods. Previous work on long-range passive RFID motion detection defined fixed rules, such as thresholding. Unlike this, we applied statistical machine-learning. We extracted relevant features from the RSSI data, and classified them using binary classifiers as moving or still. We varied the sliding window size, feature set and the classifier type to assess their joint effect on motion detection. Our experiments were performed in various scenarios including humans and multiple tags. Our results show the merits of different features and classifiers under different scenarios, which apply to other context-aware applications that involve object-based activities. We have also studied the recognition of motion type. We found that identifying motion type is more challenging than detecting object motion with current passive RFID technology.

In sections that follow, we first characterize application domains involving object-based activities (Section 3). After presenting the related work (Section 1.1), we describe our experimental setup (Section 2) and methodology for processing the RSSI data (Section 4). We report the experimental results in Section 5, and draw conclusions in Section 6.

1.1 Related Work

To date, mobility of humans and objects have been monitored using several sensing technologies. Among these, GPS [24,21] and GSM [27] are appropriate for detecting and classifying large-scale motion, such as the transportation mode of travelers (e.g. walking, driving, taking a bus). WLAN [19,12,15], active RFID [13] and accelerometers [21,2] allow for monitoring small-scale motion. However, they are not appropriate for tracking small and inexpensive objects (e.g. a toothbrush or an endotracheal tube) due to size and cost limitations. These sensors have their own energy sources, increasing not only the cost, but also the sensor size.

Passive RFID tags offer a non-intrusive and low-cost solution for tracking inexpensive and small objects. To date, wearable short-range readers have been widely preferred when detecting the passive RFID tags on these objects [25,23,6]. Although near-field technologies achieve high accuracy of interaction detection, they require humans to participate in sensing, which is intrusive in real-world applications. Even in a home setting, participants might forget to wear the readers or grasp objects with their unequipped hand [17]. These problems are more likely in stressful environments. Therefore, long-range RFID readers are required for minimizing the human intervention and building non-intrusive systems.

For detecting the movement of RFID tagged objects, RSSI data captured by long-range RFID readers can be processed in a similar way as other sensors, such as WLAN and active RFID. However, in case of passive sensors, signal strength is sensitive to several environmental factors. Polarization of the tag and reader antennas must be matched for maximum efficiency, which depends on the tag orientation. Correlation between tag-reader distance and RSSI is complicated in a realistic environment because of fading, absorption, multipath and occlusions. High tag population and human motion are significant sources of noise and occlusion in room-size environments. Depending on the application domain, the impact of these factors might vary (Table 1). Resulting RSSI can be very noisy even when both the reader and antenna are static [20,4]. We use a statistical approach to address these challenges.

Prior work on long-range passive RFID-based motion detection focused on applications such as ADLs, office activities and retail shops [10,26,7,14]. Their experiments were run in controlled settings, with few tagged objects and mostly either no or only one person present. As a result, the RSSI states in [26] were clearly distinguishable. It is unlikely to obtain such RSSI data in a trauma resuscitation setting, because it is a more challenging domain compared to ADLs, office activities and retail applications (Section 3). We study object motion detection in challenging scenarios with continuous human movement, multiple tags and concurrently moving objects. Unlike previous work that used rule-based approaches, we used statistical machine learning to process the noisy RSSI data. We report experimental results with several statistical features and classifiers, and provide guidelines on feature and classifier selection under different constraints.

2 Experimental Setup

In this section, we present our experimental setting and scenarios, designed to evaluate the effect of three main factors: 1) human presence/movement in the environment, 2) multiple tags in view 3) concurrent and nearby tag movement. These factors are highly likely in a trauma resuscitation setting and affect the radio signal strength due to interference.

3 Characterizing Applications for Object Motion Detection

The domain of application represents a unique set of properties that determine system design. In this section, we list properties related to object motion detection, identify the constraints and challenges introduced by each property, and discuss three sample applications. We set recognition of trauma tasks as our target and develop the experimental setup and methodology based on its properties (Sections 2 and 4).

Table 1 Characterization of object-based activity recognition applications

		Application Domain		
Characteristic	Requirement/Challenge	ADL	Shop./Retail	Trauma Res.
1) # of people in experimental area	Human body occludes and absorbs of radio signals. Effects become more severe with increasing number of people.	1	1	5-20
2) # of tagged objects in view	Due to the collision avoidance mechanism (Slotted Aloha Protocol [5]), number of readings from a tag decreases with the increasing number of tags in view.	5-6	1-3	10-100
3) # of concurrently moving objects	Nearby concurrently moving tags may affect the signal strength of each other and cause interference.	1-2	1-2	5-10
4) Duration of interaction	A quick interaction may not be captured as its effect on the signal strength is smoothed while windowing (Section 4.2.1).	seconds to minutes	seconds to minutes	seconds to minutes
5) Speed of object movement	Very slow movements may be perceived as if object is still, and not detected. Very fast movements may not be detected as its effect on the signal strength is smoothed while windowing.	slow (~ 0.5 m/s)	slow (~ 0.5 m/s)	slow - moderate ($0.5 - 1$ m/s)
6) Speed of human movement	Fast human movement causes frequently changing environmental characteristics, causing fluctuations even for a still tag.	slow (~ 1 m/s)	none	moderate ($\sim 1 - 3$ m/s)
7) Frequency of human movement	Frequent human movement causes frequently changing environmental characteristics and more occlusions, causing fluctuations even for a still tag.	few times per minute	n/a	many times per minute
8) Tolerance to detection latency	Less tolerance to latency restricts use of smoothing techniques	minutes	seconds to minutes	few seconds
9) Location of object interactions	Defines the region that must be covered by the RFID antennas.	scattered	scattered	scattered, mostly patient bed

Application #1: Recognizing Activities of Daily Living (ADL): ADL refers to an individual's daily self-care activities, such as making tea and brushing teeth. ADLs are monitored for assisting the elderly or disabled people, because the ability to perform ADLs is considered as an indicative for the person's functional status [25, 23, 17]. ADLs are mostly single-person activities with multiple objects. Although many objects might be in view at a time, only one to two of them can be simultaneously in motion because only one person interacts with each (Table 1).

Application #2: Surveillance in Retail/Shopping: The purpose of object monitoring here is to detect a customer's interaction with a product, and recognize what the interaction involves. The type of interaction may predict the person's goal as, e.g. examining the product or theft [14]. Typically, only the customer and multiple products are in view (Table 1).

Application #3: Recognizing Tasks During Trauma Resuscitation: Trauma resuscitation refers to the initial management of a severely injured patient in the emergency department. Recognizing tasks and providing situation-awareness during trauma resuscitation is important for supporting collaborative teamwork and improving patient care. This process requires interacting with many medical objects, most of which are uniquely associated with tasks. Many tasks therefore may be recognized based on the used medical equipment, and object motion may provide valuable information about use status (Table 1).

Trauma resuscitation requires the collaborative work of teams of clinical providers that may include more than 15 members. Tens to hundreds of tagged objects may potentially be used. Because tasks are performed in parallel, multiple objects might be in-use simultaneously. Moreover, it is a fast-paced process: speed and frequency of human movement are high. Compared to ADL and retail applications, where one person interacts with only a few objects and fewer objects are in view, trauma resuscitation is a more challenging setting.

Table 2 Interaction statistics for various items and interaction types. Average and minimum values obtained over five resuscitations.

Object	Interaction Type	Avg. # interactions	Avg. dur. (sec.)	Min. dur. (sec.)
CO_2 indicator	relocation	0.8	3	2
	use	0.4	6.5	6
Thermometer	relocation	0.5	5	2
	use	0.7	23	12
Laryngoscope	relocation	0.8	1.75	1
	use	0.8	54	26

We also performed an observational analysis on object movements during trauma resuscitation. We identified two main types of motion: linear (during relocation) and random (during use). Determining the motion type might be helpful to reduce false alarms because object relocation might be a sign, but does not necessarily indicate object use.

Duration of relocation is short and nearly constant across different objects; however duration of use varies (Table 2). Objects characterized by short interaction (<10 sec.) include wrapped items such as tubes and CO_2 indicators (Table 2). Tags on these objects are interacted with only briefly, as the object is being unwrapped. Long interactions (>10 sec.) were observed for mostly re-usable items without wrappings, such as thermometer and laryngoscope (Table 2) [22]. Because swift movement of short interacted objects cannot be detected with passive RFID, we address long-interacted objects in this work.

3.1 Environmental Setting and RFID Equipment

We performed our experiments in a room partially filled with furniture, which caused multipath fading and distortion of the RF signal (Figure 1). This setting is similar to a typical trauma resuscitation room, with a patient bed at the center, side furniture, and a free space in between. A tagged object was interacted with on the central table, and in the free space around it. We focused on this region because most object interactions happen on and around the patient bed (Table 1). Also, we do not address antenna coverage issues in this work, but evaluate the motion detection performance separately.

We used off-the-shelf RFID equipment from Alien [18]: an RFID reader (ALR-9900), three circularly polarized antennas (ALR-9611-CR) and passive tags (Squiggle ALN-9540). We used three antennas to detect movements in three dimensions (Figure 1). One antenna was ceiling-mounted (2.7 m above ground), facing the center of a plastic-top table, sized $0.76 \times 1.27 \times 0.76$ m. The other two antennas were mounted on perpendicular sidewalls 2 m above the ground (to minimize human occlusions), facing the table with an angle of 60° to the ground (slanted to include the table in the coverage area). Considering that the providers gather around the patient during trauma resuscitation, ceiling-mounted antenna is crucial for our design because it is less likely to be occluded by humans.

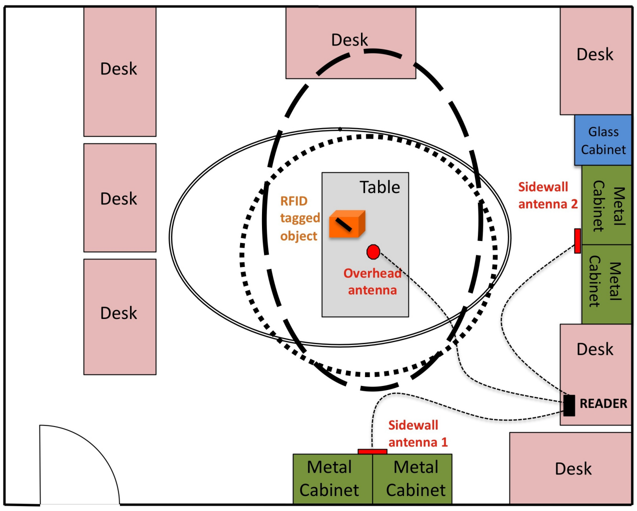

Fig. 1 Room layout and positioning of the antennas during experiments. Main coverage zones of antennas are circled. Sidewall antenna 1: dashed line; sidewall antenna 2: double solid line; overhead antenna: dotted line.

Regardless of the number of tags used in an experiment, the reader scanned for multiple tags in the environment, rather than a fast search for a single tag. Although we used a single reader, it operated in dense reader mode (DRM), which prevents interference among readers) to obtain results scalable to larger deployments with multiple readers. Also, DRM yields the best performance when tag-to-reader distance is greater than 1.5 meters [4]. Radio signal was emitted in a round robin fashion through one antenna at a time (for 0.5 seconds). The reader emitted 1 watt of RF power.

3.2 Data Collection

Our dataset consisted of 240 RSSI recording sessions. Each session lasted 60 seconds, resulting in a total of 14000 instances of motion detection. In accordance with our target application (recognizing the trauma resuscitation tasks), each session included three motion types: 1) moving linearly, 2) moving randomly, and 3) standing still (Section 3). These motion types were simulated by interacting with an RFID-tagged cardboard box ($19 \times 10 \times 7$ cm) as follows:

- Holding the box while walking in the free space with an approximate speed of 1 m/s (object moving linearly).
- Standing at the same position and wiggling the box: rotating and occluding by fingers (object moving randomly).
- Not interacting at all (object standing still)

Based on our observations of trauma resuscitation (Section 3, Table 2), we defined duration of object use as 20 seconds, and simulated random movement for 20 s. Although duration of linear motion is usually shorter than random motion, which in turn is shorter than standing still, we set their durations to 20 s to generate equal amounts of data from each motion type. The three interactions, each lasting 20 s, were performed continuously in various orders to obtain a session of 60 s.

Recordings were collected in different settings, each introducing a challenge that is likely to happen in a trauma resuscitation setting and affect the radio signal behavior:

Setting #1– Baseline: There was only one tag in view of the antennas and no movement in the environment (60 sessions). **Setting #2 – Human movement:** One tag was in view and

- Setting #2a: one person was moving (30 sessions).
- Setting #2b: two people were moving (30 sessions).

Setting #3 – Multiple tags: 10 tags were uniformly scattered and they were in view throughout the recording

- Setting #3a: there was no people movement (30 sessions).
- Setting #3b: one person was moving (30 sessions).

Setting #4 – Concurrent and nearby tag movement: 2 tags were attached to different cardboard boxes and there was no movement in the environment.

- Setting #4a: Tags moving simultaneously (30 sessions).

– Setting #4b: Second tag was moving randomly, as the target tag stood still. We studied the effect of a nearby tag movement, when the target tag was still (30 sessions).

In settings with a human presence or movement, the experimenter continuously walked around the table and from side-objects to the table with an approximate speed of 1 m/s, in accordance with a typical trauma bay (Table 1).

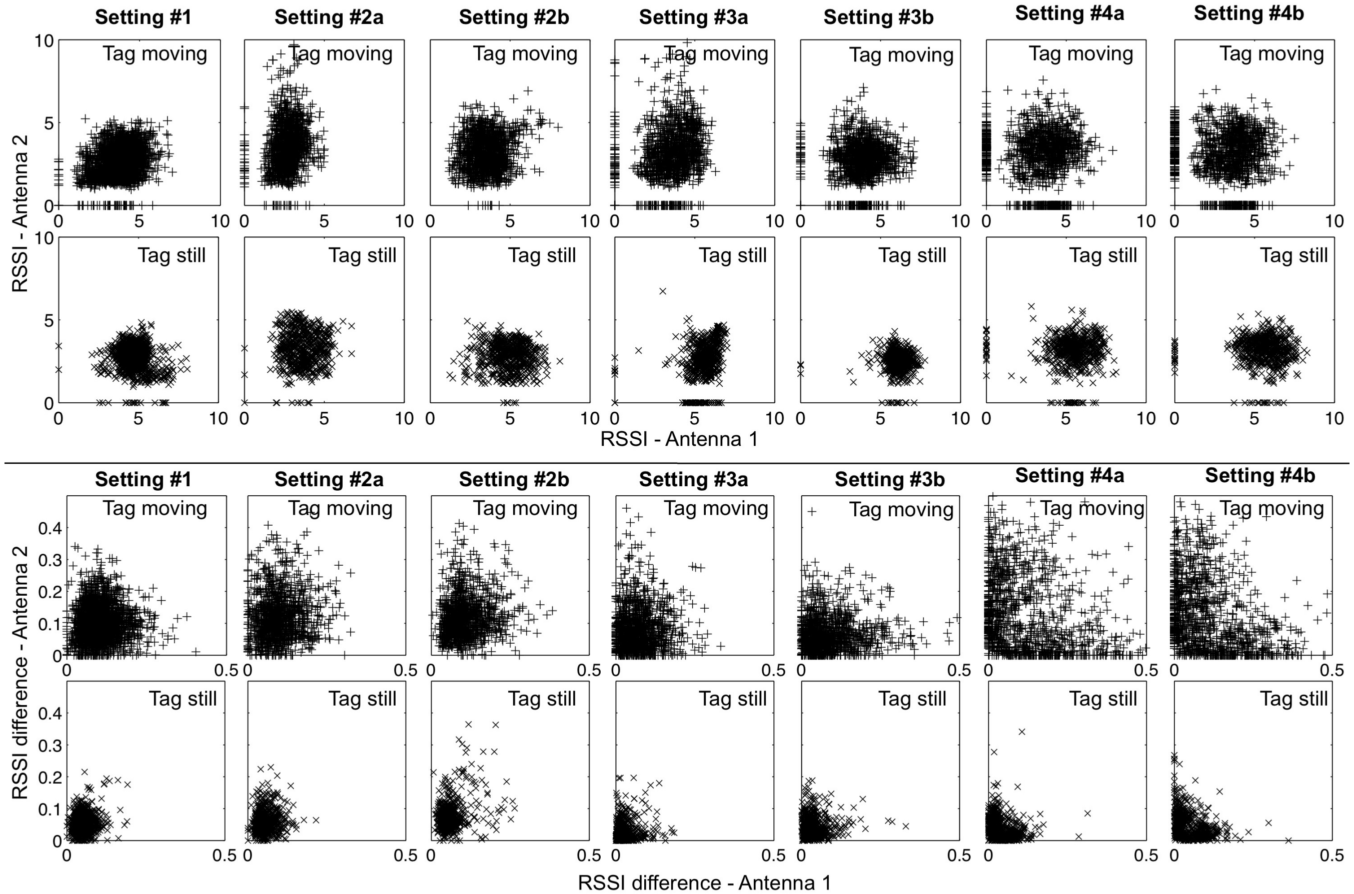

Fig. 2 RSSI values (upper two plots) and difference in RSSI of consecutive readings (lower two plots) recorded when the tag was in motion and still. Data of two antennas are shown for clarity (Antenna 1: x-axis; Antenna 2: y-axis). Environmental setting is specified at the top (Setting #1: 1 tag, no people movement; Setting #2a: 1 tag, 1 person moving; Setting #2b: 1 tag, 2 people moving; Setting #3a: 10 tags, no people movement; Setting #3b: 10 tags and 1 person moving, Setting #4a: 2 tags and no people movement (tags concurrently moving), Setting #4b: 2 tags and no people movement (nearby tag moving)).

4 Method

In this section, we first analyze the RSSI data of different motion types under different scenarios. Next, we explain our methodology for processing the RSSI data to infer the underlying high-level information about object mobility.

4.1 RSSI Data Analysis

We start by visualizing the RSSI distribution under different circumstances. Figure 2 shows the distribution of RSSI values when the tag was moving (first row) and standing still (second row). In all settings, the variance of RSSI was significantly higher when the tag was in motion. Human movement caused the RSSI to deviate, however the variance remained smaller for a still tag. In case of multiple tags, variance of RSSI is even smaller because of interpolation (explained later in Section 4.2.2). Similar to the human movement scenario (Setting #2a and #2b), concurrent and nearby tag movement caused more scattered RSSI, but variance was significantly higher when the tag was moving. These observations indicate, standard deviation is a useful feature for detecting tag motion.

We also analyzed the temporal behavior of the RSSI signal (Figure 2 – third and fourth rows). In all settings, difference of consecutive RSSI readings is much smaller when the tag is still. Compared to the deviation of RSSI values (Figure 2), difference of RSSI values provides better discrimination of moving and still states. Based on these observations, we define *standard deviation* and *difference* as our baseline features.

4.2 RSSI Data Processing

We formulate object mobility detection as a classification problem. Let X denote the RSSI received from a tag:

$$X = \{x_{t_k}^{a_k}\}_{k=1,\ldots,n}^{a_k=1,2,3} \tag{1}$$

where $x_{t_k}^{a_k}$ denotes a reading received at k'th timestamp t_k by antenna a_k. In our setup, a_k is an integer between 1 and 3.

We would like to estimate the mobility status $\hat{u}_t$ at time t $(0 < t < t_n)$, with a classification rule h:

$$\hat{u}_t = h(X) \text{ such that } \hat{u}_t = u_t \tag{2}$$

where u_t is the actual mobility status of object at time t. RSSI readings from a distant time provide little or no information about the mobility at time t. So the input can be truncated to within a window around t without affecting the performance:

$$\hat{u}_t = h(X) = h(\{x_{t_k}^{a_k}\}_{(t-w/2)<t_k<(t+w/2)}^{a_k=1,2,3}) \tag{3}$$

where w represents the window size. The length of the input to function h depends on the window size. For example, a window of 3 seconds (our typical window size) includes about 80 RSSI readings. Although this number is not too large for one object and one time instant, real-time classification for tens of objects, each generating in excess of 80 data points per decision window, can quickly become impractical. Alternatively, we can extract sufficient statistics, or features, from the RSSI data by using a function d:

$$f_t^k = d(\{x_{t_k}^{a_k}\}_{(t-w/2)<t_k<(t+w/2)}^{a_k=1,2,3}) \tag{4}$$

where f_t^k is the k'th feature coefficient at time t. Collectively, these coefficients constitute a feature vector f_t. The problem of estimating the mobility status can now be thought as learning a function g that maps the feature vector to the mobility status:

$$\hat{u}_t = g(f_t) \text{ such that } \hat{u}_t = u_t \tag{5}$$

In machine learning, the function g is known as a *classifier*, $\hat{u}_t$ a predicted label and u_t a true label. Following a supervised learning approach, we will learn function g from our labeled data (recorded during data collection). Features and labels are input to a learning algorithm which outputs a classifier.

4.2.1 Selecting the window length (w)

Several issues must be considered when choosing the window size. A larger w is useful under noisy observations as it yields smoother estimates. However short movements might be missed with a large w. The latency of the classifier is also proportional to w.

We experimented with fixed-length windows of 1.5, 2, 3, 5 and 10 sec, and adjusted the shift size $w/2$. A 1.5-sec window is selected as the shortest because our setup consists of three antennas each with a scanning time of 0.5 s. A 10-sec window is selected as the longest because the latency and smoothing of a longer window is not acceptable for our domain. The average duration of object use is 20 s, implying shorter interactions (Table 2). For detecting such interactions and generalizability of our results, we limited the window size to 10 s.

4.2.2 Interpolating the RSSI signal

In our setup, tag readings do not occur at regular intervals due to three factors:

- Because the reader is set to autonomous mode, tag readings may not arrive at regular intervals.
- Antennas are activated in a round-robin fashion to avoid interference. When the reader is transmitting from an antenna, data arrives only from that antenna.
- When multiple tags exist in the environment, read rate per tag reduces because many tags contend for wireless channel access (using the ALOHA protocol).

Before feature extraction, we preprocess the RSSI data within a window for handling the missing data intervals. First, we remove the blank intervals by concatenating the intervals when the antenna was active. Next, we interpolate the RSSI values in the window with linear interpolation. The number of interpolation points is determined based on the size of the window (such that the resulting frequency is 50 Hz).

4.2.3 Feature Extraction

When an object is standing still, its RSSI pattern may exhibit only slight deviations due to multipath fading. Changing environmental conditions (e.g. human movement or nearby tag movement) may trigger stronger deviations. However, when the object is in motion, we see a significant change in the RSSI pattern (Figure2). When detecting motion, our aim is then to capture this change in data. We have already set *standard deviation (f_1)* and *difference (f_2)* as our baseline features. We define two more features to investigate the effect of an enhanced feature set:

- *Trend (f_3):* Trend is the slope of the line fitted to the RSSI series in the current window. It captures the long-term movement of an RSSI sequence and is expected to be high for the linear movement type.
- *Delta Mean (f_4):* Delta mean represents the amount of change in average RSSI between two successive windows. Delta mean is expected to be high when the interaction status changes.

These features are computed for each antenna separately and concatenated (rather than averaging) to obtain the final feature vector. Our experiments showed that, in case of RFID, it is not sufficient to take the average of multiple receivers, because orientation is an important factor.

4.2.4 Classification

Here, we explain our methods for classifying the extracted features. Our analysis in Section 4.1 showed that, *moving* and *still* samples are overlapping in the feature space, especially in settings with human presence and movement. It is necessary to exploit the sequential properties of the data. We incorporated the sequential information in different ways. First, we applied non-temporal classification, and smoothed the output labels by removing sharp transitions between different classes. Second, we directly applied temporal classification with generative models. Finally, we used temporal classification with discriminative models, which have been previously applied to RFID-based activity recognition problems [11], but not evaluated solely on RFID data.

Non-Temporal Classification We experimented with the following non-temporal classifiers: Support Vector Machines (SVM) [1], Decision Trees (DT) [1], Random Forests [3] and Boosting [8]. The non-temporal models were built with WEKA Machine Learning Toolkit [9].

A Hidden Markov Model (HMM) based post processing was applied on the classifier output to remove spurious transitions, and smooth the label sequence. To estimate the observation probabilities of the smoothing HMM, training data is labeled and predicted labels are compared with the true labels, which were recorded during data collection.

Temporal Classification with Generative Models Determining the motion status over time is a sequential learning problem. An HMM is a generative probabilistic model, which takes the temporal (or spatial) information into consideration [1]. We built an HMM, where the actual motion states constitute the hidden state set (*still* or *moving*) and the estimated states constitute the observation symbols set (same as the hidden state set). Transition probabilities of the HMM were estimated from the training data. Observations were modeled as Gaussian mixtures with parameters estimated from the training data based on the Maximum Likelihood principle.

Temporal Classification with Discriminative Models Conditional Random Fields (CRF) are discriminative models, which specify the probability of label sequences based on the observation sequence, rather than joint probabilities. CRFs drop the independence assumption of HMMs, by allowing arbitrary relationships among the observations and dependence of hidden state probabilities on past and future observations [16]. We used the CRF-based model to train the hCRF library[1].

5 Experimental Results

In this section, we present our experimental results. Motion detection performance was evaluated with six-fold cross validation[2]. Train and test sets were determined based on sessions (Section 3.2), because readings in a session are correlated. If we include part of a session in training and the remainder in test set, this would artificially increase accuracy.

Our evaluation metric is the percentage of accurately classified labels in a session. Performance score is reported by averaging the accuracy over all sessions in the test set.

5.1 Effect of Classifier and Window Size

In this experiment, we evaluate the motion detection performance of several classifiers and the effect of window size on classification accuracy. The overlap of two consecutive windows was adjusted to be about the half-size of the window (1 second for window sizes 1.5, 2 and 3 seconds; 2 seconds for 5-second windowing; 5 seconds for 10-second windowing).

A 10-second window yields the highest accuracy for all classifiers (Figure 3). Decision Tree (DT) outperforms the other classifiers for the smallest window size (1.5 seconds). SVM outperforms these as window size increases. SVMs are known with their good generalization performance, which is also observed in our case. However a short window includes less amount of data, causing unreliable estimation of features. In this case, a simple DT yields better results. A DT is then preferred when a smaller window size is ideal, for example when detecting the motion of an object with short duration of use (e.g., CO_2 indicator in Table 2). On the other hand, SVM must be preferred when a longer window can be used, for example when detecting the motion of an object with long duration of use (e.g.,

[1] http://sourceforge.net/projects/hrcf/

[2] The dataset is randomly partitioned in 6 subsamples, and one subsample is retained as the validation set. This process is repeated 6 times, such that each each subsample is used once as validation data.

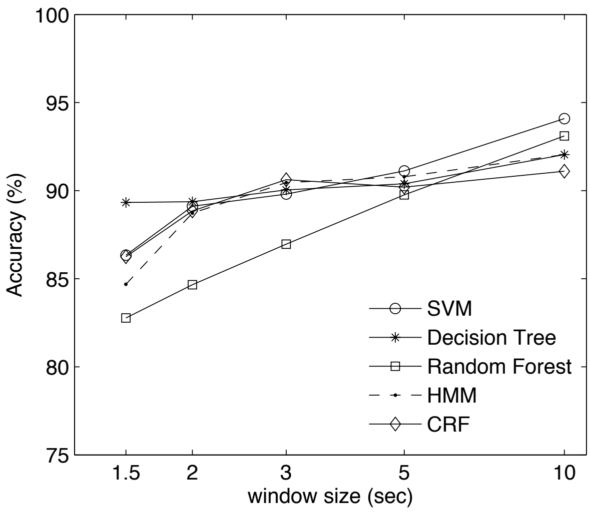

Fig. 3 Comparison of moving vs. not moving classification accuracy for different classifiers and window sizes.

laryngoscope in Table 2). The use of Random Forests did not yield any gain over DT even though it is a boosted version of DT.

Temporal classifiers, HMM and CRF slightly outperform the other classifiers in case of a three-second window, but this difference was not statistically significant. Moreover we did not observe a significant difference between HMM and CRF. Non-temporal classifiers were then better in capturing the change in RSSI signal and distinguishing these from environmental effects. It was sufficient to integrate the temporal information by smoothing the label sequence after classification.

5.2 Motion Detection Performance in Different Settings

In this section, we present the motion detection accuracy on subsets of data collected in different settings. Our aim is to investigate how motion detection performance varies from ideal to challenging conditions, which are likely to exist in dynamic settings, such as trauma resuscitation (Section 3).

Results are presented in Figure 4 for several classifiers. All of them perform best in the ideal case (Setting #1). Motion detection accuracy decreases as more human movement is introduced in the environment (Settings #2a and #2b). The reduction from Setting #1 to Setting #2a is higher than the reduction form Setting #2a to Setting #2b. As the number of people increases, the influence of the new person decreases. HMM-based classifier yielded significantly higher scores in case of human movement (Settings #2a and #2b). We concluded that, the observation model is not dramatically changed with human movement. HMMs might be useful in scenarios with a small number of tags and significant human motion.

While multiple tags (Setting #3a) decrease the accuracy in a similar way as human movement does, their combined effect is more destructive (Setting #3b).

Concurrently moving tags (Setting #4a) and movement of a nearby tag (Setting #4b) did not adversely affect the motion detection performance.

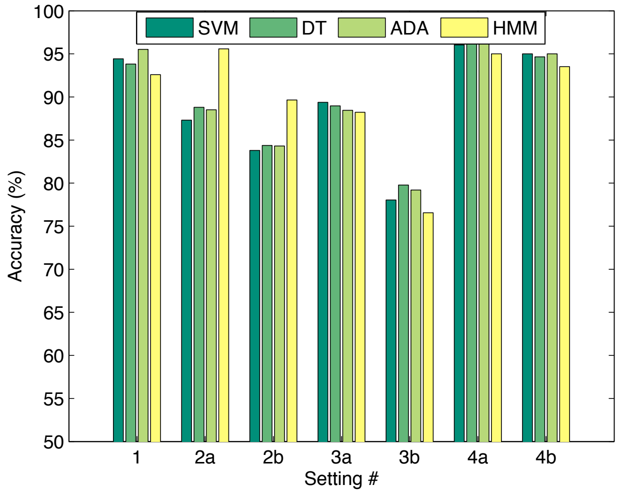

Fig. 4 Motion detection performance of several classifiers in different settings. (Setting #1: 1 tag, no people movement, #2a: 1 tag, 1 person moving, #2b: 1 tag, 2 people moving, #3a: 10 tags, no people movement; #3b: 10 tags and 1 person moving, #4a: 2 tags and no people movement (tags concurrently moving), #4b: 2 tags and no people movement (nearby tag moving)).

5.3 Effect of Interpolation

Interpolation is an extra processing for the system and may be removed if the resulting gain is negligible. In this experiment, our aim is to evaluate the effect of interpolation.

For all classifiers, the performance was significantly better when interpolation is applied (Table 3). We observed that, for the same type of motion, the difference features (f_2) computed for different settings are highly different and cannot be represented by a single parameter set. Because of this reason, decrease in accuracy is higher for HMMs: HMM is a parametric model and a single Gaussian was not able to handle all settings. Although it is possible to use a more complicated model, such as a Gaussian mixture, its training requires more data to estimate the increased number of parameters. In addition, the number of mixtures must be known in advance. Therefore, interpolation provides a simple solution to decrease the dissimilarity in case of multiple tags in the environment.

5.4 Effect of Individual Features

In this experiment, we analyze the efficiency of features described in Section 4.2.3. We included two more features, trend (f_3) and delta mean (f_4), to our baseline

Table 3 Motion detection accuracy (%) based on interpolation.

	no interpolation	with interpolation
SVM	87.0	90.2
HMM	70.9	90.5

feature set, which consisted of standard deviation (f_1) and difference (f_2). We experimented with several subsets, and present results obtained with HMM and SVM-based classifiers, as representatives of temporal and non-temporal classifiers.

A richer set of features yielded better scores for HMM (Figure 5 - left). The third feature (f_3 - trend) degrades the performance in case of short windows, but improves scores for longer windows. Estimation of this feature requires longer observation of RSSI relative to other features. As a result, its effect is accurately observable in long windows. For SVM, a richer feature set improved motion detection performance only when window size is 1.5 seconds, because the data included in this short window was not sufficient to extract reliable features. As the window size grows, added features do not have significant contribute (Figure 5 - right). We conclude that, when it is required to work with short windows and generative classifiers, such as HMMs, defining more features improves motion detection accuracy. Otherwise the baseline feature set (standard deviation and difference) performs equally well.

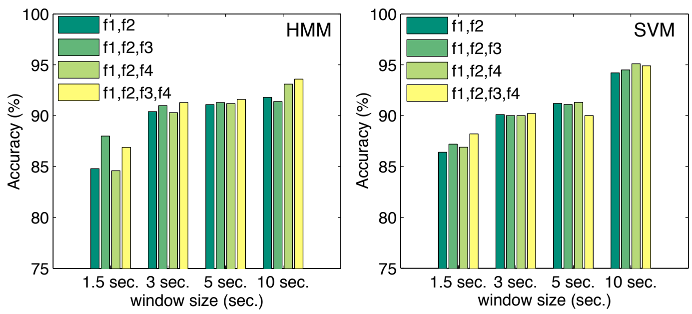

Fig. 5 Motion detection performance of different feature sets and window sizes, using an HMM-based classifier (left), and SVM-based classifier (right).

5.5 Motion Type Recognition

In this experiment, we classify the object motion as: still, moving linearly or randomly, instead of binary "still" vs. "moving". Determining the motion type might reduce false alarms during activity recognition, because random movement is a stronger indicator of object use (Section 3).

Best motion recognition scores were obtained for windows of length 3 and 5 s (Figure 6 - left). As the window size grows, linear and random movements tend to create statistically similar RSSI sequences, hence their discrimination becomes more challenging. The optimum window size can be adjusted depending on the speed of movement and distance traveled.

Using an enhanced feature set (f_1,..,f_4), accuracy rates of up to 80% are achievable with an HMM (Figure 6 - left). Compared to binary classification into still or moving (Figure 5 - left), the enhanced feature set provides greater improvement in accuracy. Therefore, recognizing the type of feature as random or linear, requires a rich set of features to be defined.

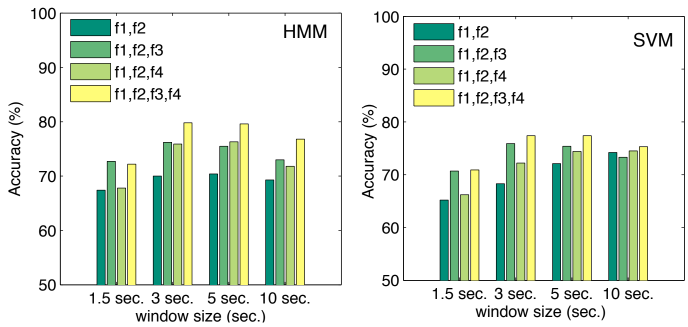

Fig. 6 Motion type classification performance of different feature sets and window sizes, using an HMM-based (left), and SVM-based classifier (right).

We also experimented with cascaded classification, where one classifier discriminates as moving and still and the second one determines the motion type as linear and random. For DT and boosting, cascaded classification yielded better scores that are close to that of SVMs (Figure 6 - right). However cascading did not provide any gain for SVM, which is originally a binary classifier, and encapsulates multiple classifiers for performing multi-class classification. Our cascading strategy was then implicitly embedded in the multi-class SVM.

6 Conclusion

We explored using passive RFID technology for long-range object motion detection. Because our target application represents a dynamic and challenging environment (due to interference caused by humans and multiple tags), we created an experimental setting that included these features and evaluated our methods on a dataset recorded in this setting. For detecting motion, we used algorithms based on statistical machine learning to process the noisy RSSI data, rather than the rule-based approaches used in previous work. We performed an extensive analysis on motion detection performance using different sliding window sizes, feature sets and classifiers. We also proposed RSSI interpolation as a solution to decrease the dissimilarity in case of multiple tags in the environment. We showed that current passive RFID technology allows for detecting object motion even in dynamic environments; however it is highly challenging to identify motion type. Our results provided an insight into what type of features and classifiers should be used under different constraints and circumstances, which are informative for other context-aware applications involving object-based activities.

References

1. Alpaydin, E.: Introduction to Machine Learning, 2nd edn. The MIT Press (2010)
2. Bao, L., Intille, S.: Activity recognition from user-annotated acceleration data. In: A. Ferscha, F. Mattern (eds.) Pervasive Computing, *Lecture Notes in Computer Science*, vol. 3001, pp. 1–17. Springer Berlin / Heidelberg (2004)
3. Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001)
4. Buettner, M., Wetherall, D.: An empirical study of UHF RFID performance. In: MobiCom'08, pp. 223–234. ACM (2008)
5. EPCglobal: radio-frequency identity protocols class-1 generation-2 uhf rfid protocol for communications at 860 mhz-960 mhz version 1.0. 9. http://www.gs1.org/gsmp/kc/epcglobal/uhfc1g2 (2004)
6. Favela, J., Tentori, M., Castro, L., Gonzalez, V., Moran, E., Martínez-García, A.: Activity recognition for context-aware hospital applications: issues and opportunities for the deployment of pervasive networks. Mobile Networks and Applictions **12**(2), 155–171 (2007)
7. Fishkin, K., Jiang, B., Philipose, M., Roy, S.: I sense a disturbance in the force: Unobtrusive detection of interactions with rfid-tagged objects. UbiComp 2004: Ubiquitous Computing pp. 268–282 (2004)
8. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: A statistical view of boosting. Annals of Statistics **28**(2), 337–407 (2000)
9. Hall, M., Frank, E., Geoffrey, H., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: An update. SIGKDD Explorations **11**(1) (2009)
10. Jiang, B., Fishkin, K.P., Roy, S., Philipose, M.: Unobtrusive long-range detection of passive rfid tag motion. IEEE Transactions on Instrumentation and Measurement **55**(1) (2006)
11. van Kasteren, T., Noulas, A., Englebienne, G., Kröse, B.: Accurate activity recognition in a home setting. In: UbiComp, pp. 1–9 (2008)
12. King, T., Kjaergaard, M.B.: Composcan: adaptive scanning for efficient concurrent communications and positioning with 802.11. In: MobiSys '08, pp. 67–80. New York, NY, USA (2008)
13. Kleisouris, K., Firner, B., Howard, R., Zhang, Y., Martin, R.P.: Detecting intra-room mobility with signal strength descriptors. In: Proc. of MobiHoc, pp. 71–80 (2010)
14. Krahnstoever, N., Rittscher, J., Tu, P., Chean, K., Tomlinson, T.: Activity recognition using visual tracking and rfid. In: Seventh IEEE Workshops on Application of Computer Vision, pp. 494 –500 (2005)
15. Krumm, J., Horvitz, E.: Locadio: Inferring motion and location from wi-fi signal strengths. In: Proc. of MobiQuitous (2004)
16. Lafferty, J.D., McCallum, A., Pereira, F.C.N.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proc. of ICML, pp. 282–289 (2001)
17. Logan, B., Healey, J., Philipose, M., Tapia, E., Intille, S.: A long-term evaluation of sensing modalities for activity recognition. In: Proc. 9th Int'l conference on Ubiquitous computing, pp. 483–500 (2007)
18. Alien Technology - Products: URL http : //www.alientechnology.com/products/index.php (last accessed Dec. 2011).
19. Muthukrishnan, K., Lijding, M., Meratnia, N., Havinga, P.: Sensing motion using spectral and spatial analysis of WLAN RSSI. In: Proc. of the 2nd European conf. on Smart sensing and context, pp. 62–76 (2007)
20. Nikitin, P., Rao, K.: Antennas and propagation in uhf rfid systems. In: IEEE International Conference on RFID, pp. 277 –288 (2008). DOI 10.1109/RFID.2008.4519368
21. Parlak, S., Jariyasunant, J., Sengupta, R.: Using smrtphones to perform transportation mode determination at the trip level. In: in Proc. of 91st Annual Meeting of Transportation Research Board (2012)
22. Parlak, S., Sarcevic, A., Marsic, I., Burd, R.: Introducing rfid technology in dynamic and time-critical medical settings: Requirements and challenges. Journal of Biomedical Informatics (2012)
23. Patterson, D., Fox, D., Kautz, H., Philipose, M.: Fine-grained activity recognition by aggregating abstract object usage. In: Wearable Computers, 2005. Proc. 9th IEEE Int'l Symp., pp. 44 – 51 (2005). DOI 10.1109/ISWC.2005.22
24. Patterson, D., Liao, L., Fox, D., Kautz, H.: Inferring high-level behavior from low-level sensors. In: UbiComp 2003: Ubiquitous Computing, pp. 73–89. Springer (2003)
25. Philipose, M., Fishkin, K., Perkowitz, M., Patterson, D., Fox, D., Kautz, H., Hahnel, D.: Inferring activities from interactions with objects. IEEE Pervasive Computing pp. 50–57 (2004)

26. Ravindranath, L., Padmanabhan, V., Agrawal, P.: Sixthsense: RFID-based enterprise intelligence. In: Proc. of the 6th Int'l conference on Mobile systems, applications, and services, pp. 253–266. ACM (2008)
27. Sohn, T., Varshavsky, A., LaMarca, A., Chen, M., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W., De Lara, E.: Mobility detection using everyday GSM traces. In: Proc. of UbiComp, pp. 212–224 (2006)
28. Turaga, P., Chellappa, R., Subrahmanian, V., Udrea, O.: Machine recognition of human activities: A survey. IEEE Transactions on Circuits and Systems for Video Technology **18**(11), 1473 –1488 (2008). DOI 10.1109/TCSVT.2008.2005594